Testimony of J. Michael Fitzmaurice, Ph.D., FAMC
on behalf of the
Agency for Healthcare Research and Quality (AHRQ)
to the
Vocabulary Task Force
of the
Health Information Technology Standards Committee
March 23, 2010

Chairman Ferguson and Co-chairman Humphreys, members of the Vocabulary Task Force, ladies and gentlemen, good afternoon. Thank you for the opportunity to provide information you requested about the United States Health Information Knowledgebase (USHIK), a metadata registry maintained by the Agency for Healthcare Research and Quality (AHRQ). AHRQ's mission is to improve the quality, safety, efficiency, and effectiveness of health care for all Americans. AHRQ has long been a strong advocate for advancing the uniformity, accuracy, and computerization of health data used for improving health research and quality of care.

I am J. Michael Fitzmaurice, Senior Science Advisor for Information Technology to the Director of AHRQ, a position I have held for 11 years. Today, I speak to you as the manager of AHRQ-USHIK, a metadata registry of the data elements that the Health and Human Services (HHS) Secretary has adopted, endorsed, or recognized for use when federal agencies exchange health information. USHIK includes a variety of metadata for health data standards, including information models, data elements, code sets, and their interrelationships for HIPAA transactions, Consolidated health Informatics, and the President's Interoperability Standards (based on HITSP's work).

History

AHRQ-USHIK follows the ISO/IEC 11179 (Parts 3 and 6) standard for metadata registries and was collaboratively designed by DoD MHS, CMS, and AHRQ. Since its inception in 1998, USHIK has had many supporting federal partners: DoD, CMS, NCI, VA, AHRQ, HHS/OS/ASPE. The fathers of USHIK are Robert Mayes, Glen Sperle, Marco Johnson, and Laura Reece. Today, I lead USHIK with my co-lead, Robert Mayes. Data Consulting Group, Inc., headquartered in Detroit, MI, is the support contractor.

How is USHIK Organized?

USHIK is actually several virtual registries linked together by common data elements. Its web site is housed at AHRQ and its data elements and metadata (data element characteristics) are freely available at ushik.ahrq.org. USHIK has several portals available and in development.

- USHIK–contains all the data elements and their metadata
- HITSP Portal–contains HITSP's data elements and their metadata, including their use cases, interoperability specifications, value sets, vocabularies, and linked to all HITSP documents

- Patient Safety Common Formats Portal (pilot)–contains AHRQ's Patient Safety Event (PSE) Common Format questions, answers (coded and uncoded) for PSE information reported and exchanged among Patient Safety Organizations and providers. Pilot projects exist on a parallel server for special access until put into production.
- Meaningful Use Measure Metadata Portal (pilot) contains example quality measures and the data elements associated with their numerators, denominators, inclusions, and exclusions.
- State All-Payer Data Metadata Portal (pilot)–contains data elements (names and other metadata) from the data entering the state's all-payer data base. This pilot will capture all-payer states' data dictionaries and relate them to HIPAA standards and each other's data definitions and representations, for comparison and harmonization purposes as well as for longitudinal maintenance.

USHIK has many information models that organize its data elements. The sub-portals begin with information models based on that portal's set of use cases. By clicking on an information model component, the user can drill down to the data elements that support that component. Of course, the user can also search for a data element by its name or by the document or standard in which it appears.

AHRQ has strongly supported HITSP for the last 4 years by supplying rapid updates to the HITSP Portal, providing technical expertise at its technical committee meetings, and supplying comparable data elements for HITSP use-case data-element choices from its stores of SDO and previous healthcare initiative items.

Vocabulary Task Force Questions
The Vocabulary Task Force has supplied questions for this panel to answer, so let me begin.

1. What vocabulary subset or value set creation and distribution services do you provide?
USHIK provides a listing of data elements, their metadata, and their links to the sources–healthcare standards from standards developing organizations (SDOs) or other authorities. Their data elements may be specified in terms of: code sets (LOINC), terminologies (SNOMED CT), classifications (ICD-9, ICD-10). These data elements are chosen by authorities that are registered under USHIK's governance structure.

- USHIK's registered authorities submit the data element value sets or vocabulary subsets they create to USHIK for inclusion.
- These authorities include ASC X12, NCPDP, HL7, ANSI HITSP, AHRQ, and others.
- USHIK relates the data elements so that users can view similar data elements and their metadata across all the content supplied by the authorities.
  - Comparison tables can be downloaded in Excel spreadsheet, XML, and pdf formats.
  - For example, you can see side-by-side in USHIK the data elements for, say, "gender" from HL7, ASC X12, CMS (CHI). [And wonder why the differences!]
  - Data element names, definitions, and their representations often differ across SDO's, hindering interoperability for the same concept.
- USHIK publishes the information received from its registered authorities at ushik.ahrq.org and primarily uses information models to guide users through the complex

relationships used in many constructs (documents).

2. Who uses your services and what is the level of use?

HITSP members and staff, researchers, vendors, and federal program managers and staff, health services researchers, standards developers, standards choosers, and others use USHIK's services.

- o We have begun capturing, organizing, and loading data element information from the most recently released HITSP documents approximately 79 documents. When published HITSP products are used, I expect USHIK's use will rise significantly.
- o USHIK freely distributes its registered data elements, their metadata, and their source links/references at ushik.ahrq.org

Specific users have commented recently:

Lynne Gilbertson, VP for Standards Development, NCPDP (March 19, 2010):

> *"For a standards development organization's perspective, the USHIK registry has been a resource to assist with researching if there are similar data elements between models. When identified, this may begin harmonization or at least a dialogue. The addition of the HITSP items has been helpful. The charts and the ease of use of the system are appreciated in this complex environment."*

John Donnelly, President, InterPro Solutions, Inc. (March 19, 2010):

> *"In short, I use this resource* [USHIK] *in discussions with my clients (physicians, nurses, and IT professionals) to help with:*
>
> *- Increasing their awareness of how data elements are to be described correctly for their use in their EHR systems*
>
> *- Understanding the value sets, i.e. content options, that are appropriate for data elements that are selected for their use for local quality and operational analytics, and which organizations "own" the upkeep of these value sets*
>
> *- Evaluating their EHR systems capabilities and support for "standards based" data elements…which then underpins focused discussions with their vendors regarding product updates*
>
> *- Understanding and facilitating data translation encoding in their interface engines across standards types, e.g. HL7, X12, etc, to maximize re-use of the data for different interface requirements without human intervention*
>
> *- Preparing them for the transition to SNOMED CT and LOINC encoding in their EHR's (and potentially the transition to ICD-10)"*

Michael Lincoln, MD, VA Medical Center in Salt Lake City, UT (March 22, 2010) stated that he uses USHIK for his HITSP projects (as Co-Chair of HITSP's Provider Technical Committee) *as well as for VA Interoperability projects* that have to be HITSP compliant, specifically:

- CHDR (Cheddar) – VA's Consolidated Health Data Repository (data transfers)
- VLER (VLER) Virtual Lifetime Electronic Record (Interoperability projects)

After finding the desired data elements in USHIK, Dr. Lincoln uses USHIK's comparison matrix feature to show the data elements' definitions, contexts, and where used. Then he traces the data elements back to their appropriate standards as well.

3. What, if any, additional services and capabilities are in active development?

Within the global registry of all its data elements, USHIK has the following major registries: HIPAA, CHI (not in active development at this time), and other data elements; HITSP, Patient Safety; State All-Payer (previously described above). All are in active development.

- Electronic Uploads–AHRQ is working with ASC X12 on a pilot to electronically upload via web services its updated information into USHIK. The development of this capability for ASC X12 could lead to all registered authorities being able to maintain their information electronically using web services and XML, with just a review by USHIK.
- New Portals–AHRQ is pilot testing portals for
  - ARRA Meaningful use measure metadata (MUMM) with a focus on quality measures. This will complement AHRQ's quality research and quality measure clearinghouse, and potentially will provide a template for other meaningful use measures.
  - State all-payer data metadata (SAPD) working with two leading states, Maine and New Hampshire, to load their data dictionaries. This will enable a user to compare the commonalities of their dictionaries' data elements. This portal can also serve as a guide for other states that are making choices of data elements' names, definitions, and representations for their all-payer data bases. Having common code sets would greatly support (1) interoperability of data for states that wish to combine their all-payer data and (2) meaningful interpretations of statistics for policy analyses and other research from each state's all-payer data.
- Specification Publishing–AHRQ is using USHIK to publish its program's technical specifications for the Patient Safety Event Common Formats, version 1.0. I expect them to be released before the end of this month. Then the Patient Safety Portal will cease to be a pilot and will become accessible to all.
- Measure Specification–AHRQ has contracted with Abt Associates, Inc., to analyze a handful of quality measures (some are also meaningful use measures) to see how their specifications and presentations might be improved from a user's perspective. That report will also be out soon.
- Nursing assessments—AHRQ has funded the University of Wisconsin at Milwaukee to develop nursing and patient safety data information regarding patient fall guidelines. Storing the metadata around the data elements in this study in USHIK will provide insight into how certain kinds of knowledge might be stored and retrieved using a metadata registry.
- Computer Query–USHIK has begun experimenting with a service oriented architecture (SOA) to investigate how a user's computer might query USHIK to obtain information, say, data element specifications for a particular quality measure—machine to machine.
- Concept Unique Identifiers (CUI's)–USHIK would like to put concept unique identifiers (CUI's) alongside each data element, code set, and terminology component. We are discussing this with Betsy Humphreys and NLM staff.
  - With CUI's, a user could find additional data elements for his/her use case that is

larger in scope than an existing but similar HITSP use case, by clicking from existing data information in USHIK into NLM's Metathesaurus. Here the user would look for additional data elements that are in the referenced vocabulary but not found in the value set of the more narrow HITSP use case that is captured in USHIK.

- o To avoid duplication, USHIK could also link to the registries of sister agencies who are willing and able to share their program data element information with the public.
  - AHRQ-USHIK has discussed with CDC the feasibility of having USHIK point to the public health information in its PHIN VADS metadata registry, rather than duplicating effort. It is mostly a matter of design and resource availability. Such efforts, however, look promising and could lead to harmonization of existing efforts, and provide a transparent and inclusive, functioning structure.

4. If applicable, what process is used to establish and revise any subsets or value sets that you distribute?

AHRQ-USHIK establishes relationships with the owners of the intellectual property (USHIK-registered authorities) that contribute their data elements and metadata to USHIK.

- o USHIK makes the information that they supply available to the public free of charge
  - The registered authorities control what is given to USHIK
- o The registered authorities update their data in USHIK as often as they desire and USHIK has the resources to accommodate updates
- o The registered authorities are the experts in their own property, make their own revisions, and establish their own value sets.
  - USHIK updates their information and undertakes version control on behalf of users
- o Current Updates: Information from ASC X12 (versions 4010 and 5010), NCPDP, and HL7 are in line for updates to the data elements and metadata they supply to USHIK, with ASC X12 being at the head of the line.

5. Based on your experience, what advice would you offer regarding best practices and pitfalls to avoid?

In developing USHIK, there are several lessons we have learned and applied:

- o Use information models based on use cases as a starting place for users to drill down to the data elements and their metadata (specifications, characteristics), such as in USHIK, Patient Safety, and MUMM Portals.
- o Trade off the value of what USHIK supplies with the cost of supplying it
  - As of July 15, for example, every HITSP-specified data element was in USHIK and linked to every HITSP document in which it was used..
    - Should we do the same linkages for the data elements in the 59+ HITSP-approved documents released at the end of January 2010?
    - Or is it sufficient to relate the data elements in those documents (say,

starting with HITSP use cases to HITSP interoperability specification to their data elements) without linking them to all the HITSP capabilities, service collaborations, content, transaction, and transaction package documents?
        - We note that the HITSP construct templates changed and some users could be confused without a USHIK road map.
- o Intellectual property has value. USHIK relies on cooperation with its registered authorities (SDO's and others) for the use of their IP
    - This can mean that in some cases we are given an SDO's data elements and their metadata but not all the information a user might need to implement the SDO's standard.
    – Don't scuttle revenue received from sales of the SDO's standards and implementation guides.
- o Work with users to anticipate their requirements and to receive their feedback
    - Save them time
    - Save them money
    - Give them accuracy
        - HITSP's Education, Communication, and Outreach (ECO) Committee is vetting the accuracy of selected information in USHIK.
        - On each page of USHIK is a query button for asking questions and making user comments.
- o Recognize that a metadata registry is more than an ordered, related collection of data elements and their metadata across use cases.
    - Governance–USHIK has registered authorities and their points of contact, as well as provisions in place for self-governance of each registry by the owner
    - Versioning–USHIK has the capability to show which data elements were valid for a particular use case at a particular time, and which data elements are valid for today's uses
    - Intellectual Property–USHIK has agreements with the authorities that supply USHIK's content where that content is not freely available
    - Services–USHIK is working on automated data loading from its authorities for the near future. Further out, USHIK will develop automated ways of searching and retrieving its data elements and their metadata.

Recommendations

- o Develop national use cases that are clinically and administratively valid, with requests for detailed data element specifications
- o Specify in great detail the standards and data elements needed to fulfill the functions of the use case
    - Choose first from existing standards of SDO's
- o When there are gaps and overlaps,
    - Charge the SDO's to fill those gaps
    - Charge the SDO's to eliminate those gaps (harmonization first, if needed)
- o Designate a metadata registry (preferably USHIK) to receive the recommended data elements and their metadata for those use cases
    - When it is inconvenient or impractical (intellectual property considerations, for

example) to specify each data elements or their metadata (e.g., complete code sets), specify the terminologies for them and provide working examples in the registry. Define reusable conceptual domains for the registry.
    - Link the use-case, data-element information to their names, definitions, and representations that are found in the SDO's standard from which they come–if the SDO's data element information has been altered by the registered authority.
    - Link those data elements to the terminologies to which they belong, so that the user can obtain appropriate data elements for similar but larger use cases than originally specified
- Consider linkages from use cases to a registry of their required data elements (when officially required by, say, a biosurveillance use case, or an immunization use case) to the vocabularies stored at the National Library of Medicine and the National Cancer Institute
    - Consider linking the designated metadata registry to CDC's PHIN VADS registry for public health value sets of data elements that CDC maintains
    - Consider linking to CDISC's metadata registry for clinical trial data elements if the data element information need is not fully specified in Clinical Research Component of HITSP over time
- Partner with users. Veterans Affairs is a valued partner and has supplied USHIK with the expertise to rapidly load the data elements for HITSP's capabilities and service collaborations in 2009 and HITSP's recently released standards documents, including quality measures, in January 2010.
    - With VA's help, we met the deadline of having every HITSP-specified data element properly linked to all of HITSP's documents as of July 15, 2009.
    - Also with VA's help in 2006, USHIK piloted XML exchange of HL7 information between the VA registry and USHIK.
    - Federal partners of USHIK have included: AHRQ, CMS, VA, NCI, DoD, and HHS OS/ASPE.
    - Oversight of USHIK has been provided by the ANSI Health Informatics Standards Board (HISB), ANSI HITSP, Federal Health Architecture, and USHIK's partners.
- Find a way to shorten the process of updating HIPAA and meaningful use regulations.

Conclusion

Users of health care data standards want a "one-stop shop", a single source of information about the data elements, their metadata, and the standards from which they come. Reducing the effort it takes to move from an implementation specification to its data elements and metadata is essential for efficient implementation of standards. If one trusted expert can develop the path to get to that endpoint (data elements names, definitions, representations), why not have that expert share the path in a metadata registry. We don't need 10,000 people to invent or even to travel that path. One good expert can take a picture of the endpoint and place it in a metadata registry for all to use.

In conclusion, AHRQ-USHIK can be that one-stop shop. AHRQ-USHIK welcomes additional partners to help maintain its accuracy and helpfulness. AHRQ-USHIK also welcomes constructive criticism about its contents and format for searching and displaying its contents.

The more uniform, accurate, and computerized health data are, the more useful and robust will be AHRQ-funded research findings based on these data, and the more useful will be the tools developed from AHRQ's research.